

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

528 708

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
8. April 2004 (08.04.2004)

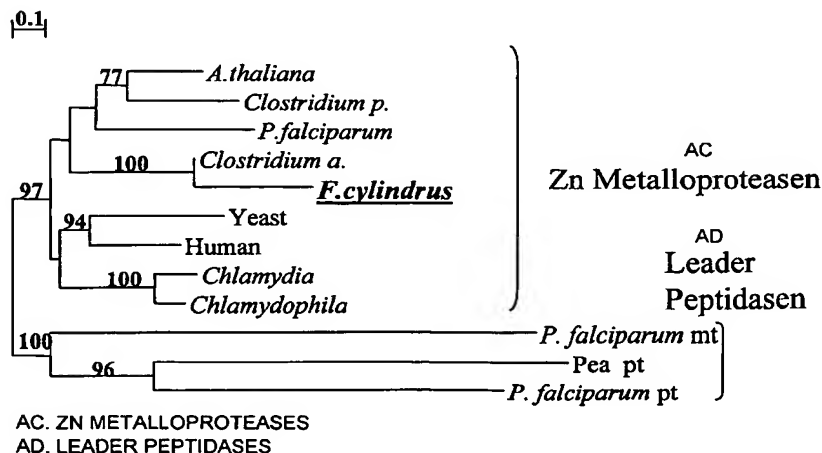
PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/029253 A2

- (51) Internationale Patentklassifikation⁷: C12N 15/57, 9/50, A61K 38/48
- (21) Internationales Aktenzeichen: PCT/DE2003/003180
- (22) Internationales Anmeldedatum:
21. September 2003 (21.09.2003)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:
102 44 842.6 22. September 2002 (22.09.2002) DE
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): STIFTUNG ALFRED-WEGENER-INSTITUT FÜR POLAR- UND MEERESFORSCHUNG [DE/DE]; Columbusstrasse, 27568 Bremerhaven (DE).
- (72) Erfinder; und
(75) Erfinder/Anmelder (nur für US): MOCK, Thomas [DE/DE]; Bürgermeister-Smidt-Strasse 178, 27568 Bremerhaven (DE). VALENTIN, Klaus [DE/DE]; Hafenstrasse 62, 27576 Bremerhaven (DE).
- (81) Bestimmungsstaat (national): US.
- (84) Bestimmungsstaaten (regional): europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR).
- Veröffentlicht:
— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts
- Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: NUCLEIC ACID SEQUENCES CODING FOR PROTEOLYTIC ENZYMES IN THE FORM OF SPECIFIC PROTEASES, CORRESPONDING POLYPEPTIDES AND USE OF THE SAME

(54) Bezeichnung: FÜR PROTEINSPALTENDE ENZYME IN FORM VON SPEZIFISCHEN PROTEASEN KODIERENDE NUKLEINSÄURESEQUENZEN, ZUGEHÖRIGE POLYPEPTIDE UND VERWENDUNG VON ALLEN



(57) Abstract: The invention relates to enzymes from the specific groups of calpain proteases, which take part in a plurality of metabolic processes, e.g. by acting on the apoptosis, for certain cancerous diseases and during cell migration. The invention also relates to enzymes from the group of metalloproteases which develop, e.g. activities during fertilisation, and take part in blood pressure regulation as ACEs and as collagenases in collagen metabolism. A potential need of such enzymes is mainly covered today by plant and animal sources. Known organisms in which such proteases naturally occur exclusively come from warmer regions, such that a costly heat supply necessitating complex apparatus is required during the production thereof. Organisms comprising enzymes adapted to the cold are known however. The invention thus also relates to nucleic acid sequences coding for a calpain 7 protease and a zinc metalloprotease, which originate from the marine diatoms *Fragilariopsis cylindrus* adapted to the cold and correspond to SEQ ID No.1, SEQ ID No.2 or functional variants or fragments having at least 8 nucleotides thereof.

[Fortsetzung auf der nächsten Seite]

WO 2004/029253 A2



(57) **Zusammenfassung:** Enzyme aus der spezifischen Gruppe der Calpain-Proteasen sind an vielen Stoffwechselvorgängen beteiligt, z.B. durch Einflussnahme auf die Apoptose, bei bestimmten Krebserkrankungen und bei der Zell-Migration. Enzyme aus der Gruppe der Metalloproteasen entwickeln z.B. Aktivitäten bei der Befruchtung, sind als ACE bei der Blutdruckregulierung und als Collagenasen beim Collagen Stoffwechsel beteiligt. Ein etwaiger Bedarf an solchen Enzymen muss heute hauptsächlich aus pflanzlichen und tierischen Quellen gedeckt werden. Bekannte Organismen, in denen solche Proteasen natürlicherweise vorkommen, stammen ausschließlich aus wärmeren Regionen, sodass bei ihrer Produktion eine ökonomisch und apparatetechnisch aufwändige Wärmezufuhr erforderlich ist. Bekannt sind jedoch auch Organismen, die kälteangepasste Enzyme aufweisen. Die Erfindung beansprucht daher für eine Calpain-7-Protease und eine Zink-Metalloprotease kodierende Nukleinsäuresequenzen, die aus der kälteangepassten marinen Diatomee *Fragilariopsis cylindrus* stammen und gemäß SEQ ID No.1, SEQ ID No.2 oder als funktionelle Varianten oder als Abschnitte mit mindestens 8 Nukleotiden davon ausgebildet sind.

Für proteinspaltende Enzyme in Form von spezifischen Proteasen kodierende Nukleinsäuresequenzen, zugehörige Polypeptide und Verwendung von allen.

5 **Beschreibung**

Die Erfindung bezieht sich auf für proteinspaltende Enzyme in Form von speziellen Proteasen kodierende Nukleinsäuresequenzen, auf die zugehörigen Polypeptide und auf Verwendungen sowohl der Nukleinsäuresequenzen selbst
10 als auch der zugehörigen Polypeptide.

Proteasen (auch: Proteinasen) sind eiweißspaltende Enzyme („proteolytische Hydrolasen“) die nach ihren Wirkungsmechanismen unterschieden werden. Eiweiße sind typisch gefaltete Polypeptidketten. Die Proteasen werden daher
15 auch als Peptidasen bezeichnet und in Endo- und Exopeptidasen unterschieden. Endopeptidasen spalten die Peptidketten im Inneren auf und erzeugen so kleinere Teilstücke, Exopeptidasen können die endständigen Aminosäuren abspalten. Proteolytisch wirkende Enzyme sind in biologischen Organismen allgegenwärtig und erfüllen typspezifische Zwecke. Sie können
20 daher einerseits funktionsbezogen zu Therapiezwecken eingesetzt werden oder technischen Reinigungszwecken gegen proteinhaltige Verschmutzungen dienen. Die Familien der Calpain-Proteasen (CANP) und Metalloproteasen (MP) sind wichtige Vertreter der Proteasen und Gegenstand intensiver Forschungen.

25

Calpain-7-Protease:

Bisher sind ca. 15 verschiedene CANP bekannt. Die am weitesten erforschten CAPN bestehen aus einer großen 80kDa (kilo Dalton) - und einer kleinen 30kDa - Untereinheit. Die große Untereinheit besteht ihrerseits aus vier und
30 die kleine Untereinheit aus zwei funktionalen Domänen. Die Domäne IV der großen und die Domäne VI der kleinen Untereinheit sind regelmäßig über ihre Faltungsformen („EF-hands“) als calziumbindende Domäne ausgeprägt. Es

wird vermutet, dass die proteolytische Aktivität dieser CAPN erst durch die gebundenen Calziumatome ausgelöst wird. Die neueren CAPN (5,6,7,10,13; vergleiche z.B. Veröffentlichung I) verfügen dagegen in der Domäne IV der großen Untereinheit über keine solchen, calziumbindenden Strukturen.

5 Grundsätzlich sind alle Mitglieder der CAPN-Familie gewebsspezifisch ausgebildet und erfüllen so die unterschiedlichsten Aufgaben.

CAPN nehmen oft Schlüsselstellungen in Stoffwechselwegen ein und sind am Verlauf von Krankheiten beteiligt. Z.B. sind die CAPN 1 und 2 Komponenten

10 der Reaktionskaskade der Apoptose („programmierter Zelltod“) und damit am Verlauf der Alzheimer-Krankheit beteiligt. Auch auf Krebserkrankungen können CAPN Einfluss nehmen. Dabei werden z.B. Brust- und Darmkrebs durch die Erhöhung der zellulären Konzentration des p53-Markerproteins ausgelöst. Der Konzentrationsanstieg wird durch die Inhibierung von beteiligten CAPN erreicht,

15 die die p53-Konzentration in gesunden Zellen normalerweise sehr gering halten (vergleiche **Veröffentlichung I**: „The calpain family and human disease“, Yuanhui Huang und Kevin K.W.Wang in Trends in Molecular Medicine Vol.7 No.8 August 2001, pp 355-362).

20 Calpaine spielen eine besondere Rolle bei der Zell-Migration in der extrazellulären Matrix. Zellen wandern, indem sie sich am hinteren Ende auflösen und am vorderen Ende aufgebaut werden. Calpaine sind dabei durch proteolytische Spaltung von Proteinkomplexen des Stützapparats am Ende der Zelle an dieser Wanderungsbewegung beteiligt (vergleiche **Veröffentlichung II**:

25 „V-SRC's hold over actin and cell adhesions“, Frame, Fincham, Carragher, Eyke in Nature Reviews, Molecular Cell Biology, Vol.3, April 2002, pp 233-245).

Die beschriebenen Erkenntnisse wurden durch Untersuchungen an Wirbellosen, Säugetieren und dem Menschen gewonnen. CAPN in Pflanzen

30 und Pilzen sind bisher nur durch Genom- und EST-Projekte identifiziert worden. Bei Pflanzen wird eine Wirkung in der Pathogenabwehr, bei Pilzen bei der

Anpassung an alkalische Lebensbedingungen vermutet. Bei der gesamten Klasse der Kieselalgen sind CAPN bisher nicht bekannt.

Enzyme aus der Familie der Calpaine (jedoch nicht Calpain-7) werden in der
5 Patentliteratur z.B. in **EP1214427**, **CA2321270** und **DE10031932** offenbart. In Bezug auf die vorliegende Erfindung handelt es sich bei diesen Veröffentlichungen um technisch-wissenschaftlichen Hintergrund.

Zink-Metalloprotease:

10 Metalloproteasen (MP) sind unter anderem regelmäßig in den Mitgliedern der sogenannten ADAM-Familie (a disintegrin and metalloprotease) der Transmembran-Proteine enthalten, die aus einer Desintegrin- und einer MP-Domäne bestehen und Zelladhäsions- und Proteaseaktivität im Zusammenhang mit Vorgängen bei der Befruchtung, der Entstehung von
15 Nervengewebe und bei Entzündungsreaktionen entfalten (vergleiche **Veröffentlichung III**: „Autotrope Signaltransduktion durch membranständigen Tumor-Nekrose-Faktor TNF“, Doktorarbeit Uni Stuttgart, E.Haas, 1999, pp 6-8).

Ferner spielen Metalloproteasen eine wichtige Rolle im Renin-Angiotensin-
20 Aldosteron-System (RAS). Die Aspartyl-Protease Renin spaltet Angiotensinogen in Angiotensin I. Das Angiotensin-Converting-Enzym (ACE), das anschließend Angiotensin I in Angiotensin II spaltet, ist eine Zink-Metalloprotease (ZnMP). Angiotensin II hat eine gefäßverengende und damit Blutdruck steigernde Wirkung und fördert die Freisetzung von Aldosteron in der
25 Nebennierenrinde woraus letztlich Hypertrophie am Herzen folgt. ACE ist also ein wichtiger Stoff zur Regulierung der Kreislauffunktionen und muss bei bestimmten Herz-Kreislaufferkrankungen unterdrückt werden (ACE-Inhibitor) (vergleiche **Veröffentlichung IV**: Cardiovascular Physiology Concepts, University of Ohio, R.E.Klabunde,v.04.06.2002 aus <http://www.oucom.ohiou.edu/CVPhysiology/BP015.htm> (Stand 20.08.2002)).
30

Weiterhin sind Metalloproteasen als Collagenasen am Abbau von Stützgewebe im Körper beteiligt. Collagene sind Gerüsteiweißkörper und bilden als gegen enzymatische Angriffe schützender Hauptbestandteil interzellulärer Stützsubstanz 25% des Körpergewebes. Durch Einlagerung mineralischer Kristalle können Collagene zu harten Knochen- und Zahnsubstanzen werden. Im Collagen-Lebenszyklus dienen Collagenasen zur Verarbeitung nicht mehr benötigter Collagenanteile. Collagenasen spielen eine Rolle bei der Chemonukleolyse, einer Behandlungsmethode bei Bandscheibenschäden durch Auflösung des Bandscheibenkerns (nucleus pulposus).

10

Enzyme aus der Familie der Metalloproteasen werden in der Patentliteratur z.B. in **WO9964610** (Zink-Metalloprotease), **US2002068055** (Metallo-Endopeptidase) und **US5750391** (Metallo-Endopeptidase) offenbart. In Bezug auf die vorliegende Erfindung handelt es sich bei diesen Veröffentlichungen auch um technisch-wissenschaftlichen Hintergrund.

15

Aus der **Veröffentlichung V** „Zahllose Geheimnisse der Natur“ von K. Eske (vgl. BioLOG, 3.Ausgabe, Februar 2000, pp 2/3, abrufbar unter <http://www.bioregio.org/BioLOG-3.pdf>, - Groß-/Kleinschreibung bei Seitenaufruf im Internet beachten! - Stand 01.09.2002) ist es bekannt, kälteangepasste Enzyme aus Bakterien, die in Tiefseeregionen vorkommen, zu isolieren. Neben dem Vorteil, dass kälteangepasste Enzyme zur Expression keine erhöhten Temperaturen benötigen, ergibt sich bei den entsprechenden Organismen aus den Tiefseeregionen ein großes Vorkommenspotential, da 80% der Erdoberfläche bedeckenden Wassers eine Temperatur unter 5 °C aufweist. Die Temperatur des Aktivitätsmaximums dieser Organismen und ihrer funktionellen Bestandteile liegt deutlich unter dem anderer Organismen aus temperierten und tropischen Breiten. Durch den Einsatz von kälteangepassten Mikroorganismen, auch in gentechnisch modifizierter Form, ist damit eine Produktion unter niedrigen Temperaturen mit einem gegenüber den herkömmlichen Gewinnungsverfahren mit nicht kälteangepassten Organismen

20

25

30

deutlich geringerem Energieeinsatz entscheidend wirtschaftlicher als bei deren sonst üblichen Aktivitätstemperaturen.

Vor dem Hintergrund dieser bedeutenden Erkenntnisse ist es daher die
5 **Aufgabe** für die vorliegende Erfindung, zur wirtschaftlichen Produktion der erfindungsgemäßen Proteasen einen kälteangepassten Organismus zu finden, der Nukleinsäuresequenzen aufweist, die für kälteangepasste Enzyme in Form von spezifischen Proteasen bei niedrigen Prozesstemperaturen (um 0°C) kodieren. Die erfindungsgemäße **Lösung** hierfür ist dem Anspruch 1 zu
10 entnehmen. Es werden zwei verschiedene Nukleinsäuresequenzen aus der kälteangepassten Diatomee *Fragilariopsis cylindrus* beansprucht, von denen die eine für eine Calpain-7-Protease und die andere für eine Zink-Metalloprotease kodiert. Vorteilhafte Weiterbildungen und Anwendungen, die sich auch auf die zu den beanspruchten Nukleinsäuresequenzen zugehörigen
15 Polypeptide beziehen, sind den unter- und nebengeordneten Ansprüchen zu entnehmen.

Mit der erfindungsgemäßen Lösung werden die an die Erfindung gestellten Anforderungen optimal erfüllt. Zum einen weist der die beanspruchten
20 Nukleinsäuresequenzen aufweisende Organismus spezifische Proteasen in Form sowohl eines Calpain-7-Protease-Enzyms als auch eines Zink-Metalloprotease-Enzyms auf, zum anderen entstammt der Organismus der antarktischen See, sodass diese Enzyme zusätzlich kälteangepasst sind und zu ihrer Aktivität keine Wärmezufuhr benötigt wird. Die Kenntnis solcher, die
25 beanspruchten Nukleinsäuresequenzen enthaltenden Gene ist von elementarer Wichtigkeit, wenn es z.B. um die Bereitstellung großer Mengen solcher Nukleinsäuresequenzen bei geringem Energieaufwand geht. Die mikrobielle Synthese dieser Stoffe erzielt mit der speziellen, schnell wachsenden kälteangepassten Kieselalge *Fragilariopsis cylindrus* unter niedrigen
30 Temperaturen hohe Erträge.

Durchgeführte Sequenzanalysen bestätigen, dass die erfindungsgemäß beanspruchten Nukleinsäuresequenzen für eine Calpain-7-Protease bzw. eine Zink-Metalloprotease kodieren. Durch die heutigen, entwickelten Methoden der automatisierten Sequenzierung von Nukleinsäureabschnitten ist es möglich geworden, in überschaubaren Zeiträumen gezielt Gene mit gewünschten Eigenschaften aus aussichtsreichen Organismen zu isolieren. Umfangreiche öffentliche Datenbanken mit bereits bestimmten, bekannten Funktionen zugeordneten Sequenzen dienen der schnelleren Verifizierung der Ergebnisse der mühsamen Suche. Zur Bereitstellung des für die Sequenzierung aufbereiteten Grundmaterials aus der speziellen Kieselalge dient eine Reihe von an sich bekannten Arbeitsschritten:

1) Isolation und Kultivierung des Organismus *Fragilariopsis cylindrus*

Isolation : Während einer Antarktisfahrt mit dem deutschen Forschungseisbrecher „Polarstern“ in die Weddell-See wurde die Kieselalge *Fragilariopsis cylindrus* aus dem Meereis isoliert. Die Artbestimmung erfolgte in einfacher Weise durch die Typisierung der Struktur der Schale (vergleiche **Veröffentlichung VI** von Medlin & Priddle : „Polar marine diatoms“, 2nd edition, British Antarctic Survey, Cambridge 1990, pp 182, 192).

Kultivierung : Die Kieselalge wurde bei 0°C in einem nährsalzangereicherten Medium 2 x f/2 bei 10 µmol Photonen m⁻²s⁻¹ mit 24h Licht, gehalten (vergleiche: **Veröffentlichung VII** von Guillard & Ryther : “Studies of marine plancton diatoms, I. *Cyclotella nana* (Husted) and *Detonula confervacea* (Cleve)“, 1962, Can.J.Microbiol. 8, pp 229:239). Für eine gesteigerte Expression der für die Kälteanpassung der Art verantwortlichen Gene wurden die Algen in der Mitte der exponentiellen Wachstumsphase auf -2°C abgekühlt. Dies entspricht dem Gefrierpunkt von Meerwasser. Nach 5 Tagen wurden die Boten-RNA (mRNA, messenger RNA) aus den Algen isoliert. Die mRNA repräsentieren alle gerade

aktiven Gene, also auch diejenigen, die für die Kälteanpassung verantwortlich sind.

2) Isolation der mRNA

5

Die gesamte RNA wurde mit dem RNAeasy Plant Mini Kit (Firma Qiagen) isoliert. Aus ca. 100 µg RNA konnte mit Hilfe des Oligotex mRNA Midi Kit (Firma Qiagen) ca. 800 ng RNA für die cDNA-Synthese isoliert werden.

10

3) Herstellung und Screening einer cDNA-Bank

Die cDNA-Bank wurde auf der Basis der mRNA mit dem SMART cDNA-Library Construction Kit (Firma Clontech) hergestellt:

15

A) Von der mRNA wurde dazu mit Hilfe von Oligonukleotiden und CDS III/3' Primern der erste cDNA-Strang synthetisiert.

B) Anschließend wurde die Doppelstrangsynthese mit Hilfe der LD-PCR (long distance polymerase chain reaction) im Eppendorf-Thermocycler unter folgendem Programm realisiert:

20

C)

1. 5 min Denaturierung bei 95°C, anschließend 20 Zyklen von 6 min bei 68°C und 2 min bei 95°C.

25

2. Nach einem Sfil-Verdau (mit Restriktionsenzym aus *Streptomyces fimbriatus*) der cDNA wurde sie in CHROMA Spin-400 Säulen der Größe nach fraktioniert, so dass nur cDNAs der Länge >400bp (Basenpaare) für die Klonierung zum Einsatz kamen.

30

3. Diese cDNAs wurden über Nacht bei 16°C in TriplEX2 Vektoren ligiert, die von λ-Phagen aufgenommen werden konnten. Der Titer dieser cDNA-Bank lag bei 2.7×10^9 pfu/ml (plaque forming units / ml).

4. Ein Blau-Weiß-Screening mit IPTG (Isopropyl-b-D-Thiogalactosid) und X-Gal (X-Galactosid) zeigte eine Rekombinationseffizienz von 70%.

4) Sequenzanalyse

5

Positive Phagen-Plaques wurden vom 5'-Ende mit λ -Primern vom Qiagen-Sequencing-Service sequenziert. Die Sequenzen wurden in der Genbank bei NCBI (National Center for Biotechnology Information) mit der BLASTX-Option auf ihre Homologien zu vorhandenen Sequenzen untersucht (BLAST-Protokoll vom 14.12.2001, Calpain-7-Protease; BLAST-Protokoll vom 24.04.2002, Zink-Metalloprotease). Für die Calpain-7-Protease konnte aus ca. 400 Sequenzen eine homologe Sequenz, für die Zink-Metalloprotease mindestens eine homologe Sequenz gefunden werden.

15 In **Figur 1** ist das Ergebnis einer phylogenetischen Analyse für die Calpain-7-Protease aus der Diatomee *Fragilariopsis cylindrus* dargestellt. Sie gruppiert mit signifikantem bootstrap support (allgemein bekannte mathematisch-statistische Methode) von 99% mit anderen Calpain-7-Proteasen. In **Figur 2** ist das Ergebnis einer phylogenetischen Analyse für die Zink-Metalloprotease aus
20 *Fragilariopsis cylindrus* dargestellt. Sie gruppiert mit signifikantem bootstrap support von 99% mit anderen Zink-Metalloproteasen.

Es handelt sich daher bei den kälteangepassten Enzymen, die von den beanspruchten Nukleinsäuresequenzen nach der Erfindung aus der Diatomee
25 *Fragilariopsis cylindrus* kodiert werden, mit einer Sicherheit von 99% ebenfalls um eine Calpain-7-Protease bzw. um eine Zink-Metalloprotease.

Patentansprüche

1. Für proteinspaltende Enzyme in Form von speziellen Proteasen kodierende
5 Nukleinsäuresequenzen,
dadurch gekennzeichnet, dass
die Nukleinsäuresequenzen aus der kälteangepassten marinen Diatomee
Fragilariopsis cylindrus stammen und für eine Calpain-7-Protease gemäß
SEQ ID No.1 und für eine Zink-Metalloprotease gemäß SEQ ID No.2 oder für
10 funktionelle Varianten beider Proteasen kodieren oder als Abschnitte mit
mindestens 8 Nukleotiden davon ausgebildet sind.
2. Nukleinsäuresequenzen nach Anspruch 1,
dadurch gekennzeichnet, dass
15 die Nukleinsäuresequenzen als DNA oder RNA, vorzugsweise als
doppelsträngige DNA ausgebildet sind.
3. Nukleinsäuresequenzen nach Anspruch 1 oder 2
dadurch gekennzeichnet, dass
20 die Nukleinsäuresequenzen in Vektoren, vorzugsweise in Expressionsvektoren
enthalten sind.
4. Verwendung der Nukleinsäuresequenzen nach Anspruch 3 zur Expression
oder Überexpression der Enzyme Calpain-7-Protease und/oder Zink-
25 Metalloprotease in Wirtsorganismen.
5. Zu den Nukleinsäuresequenzen gemäß Anspruch 1 oder 2 gehörige
Polypeptide, die aus mit den Nukleinsäuresequenzen gemäß SEQ ID No.1 und
SEQ ID No.2 kodierten Aminosäuresequenzen, als funktionelle Varianten
30 davon oder als Abschnitte mit mindestens 6 Aminosäuren davon bestehen.

6. Verwendung der Enzyme Calpain-7-Protease und Zink-Metalloprotease nach Anspruch 1 zu Therapie Zwecken
7. Verwendung der Enzyme Calpain-7-Protease und Zink-Metalloprotease nach Anspruch 1 zu Reinigungszwecken bei proteinhaltigen Verschmutzungen.
8. Verwendung der Polypeptide nach Anspruch 5 zu Therapie Zwecken.
9. Verwendung der Polypeptide nach Anspruch 5 zu Reinigungszwecken bei proteinhaltigen Verschmutzungen.

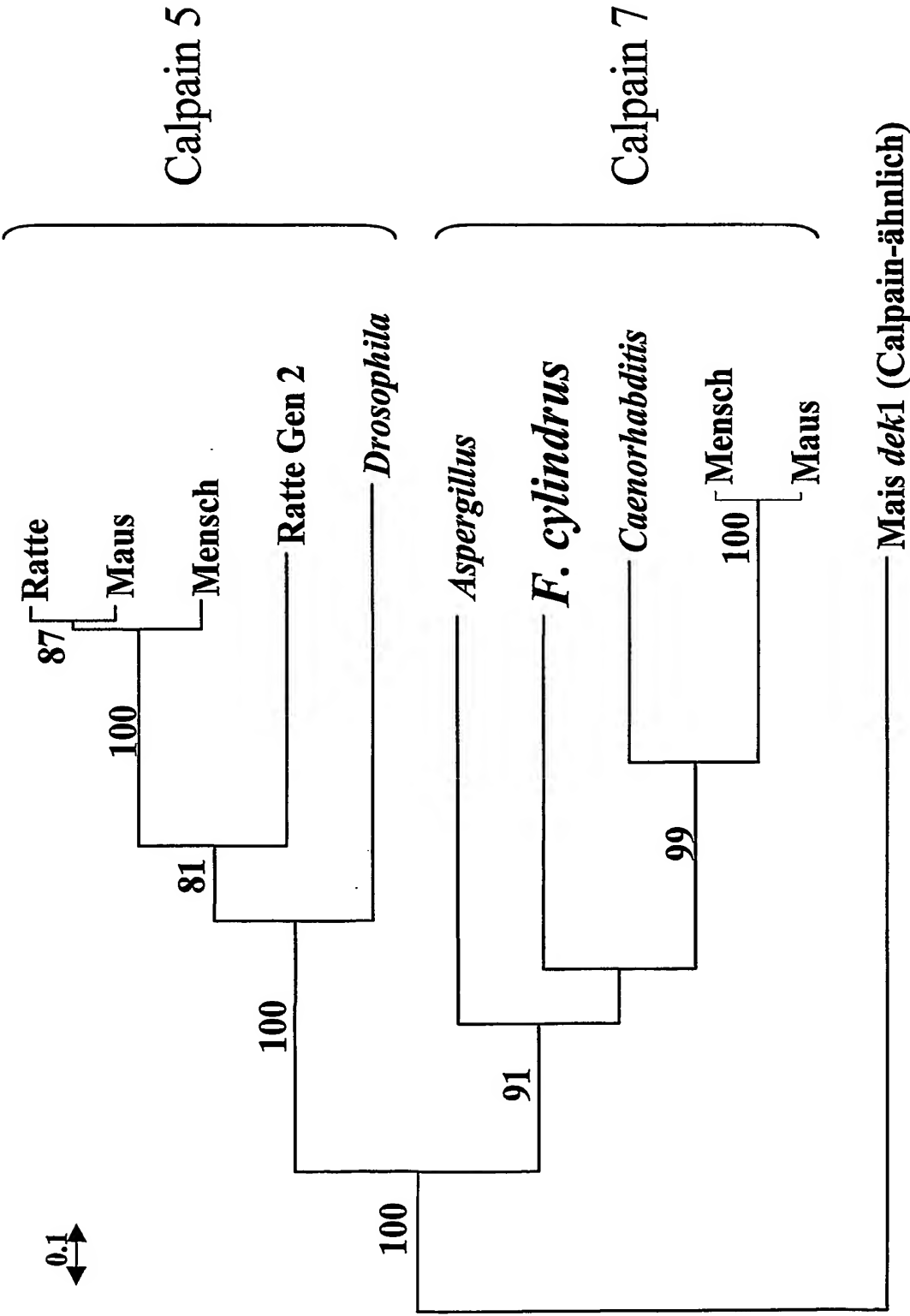


Fig. 1

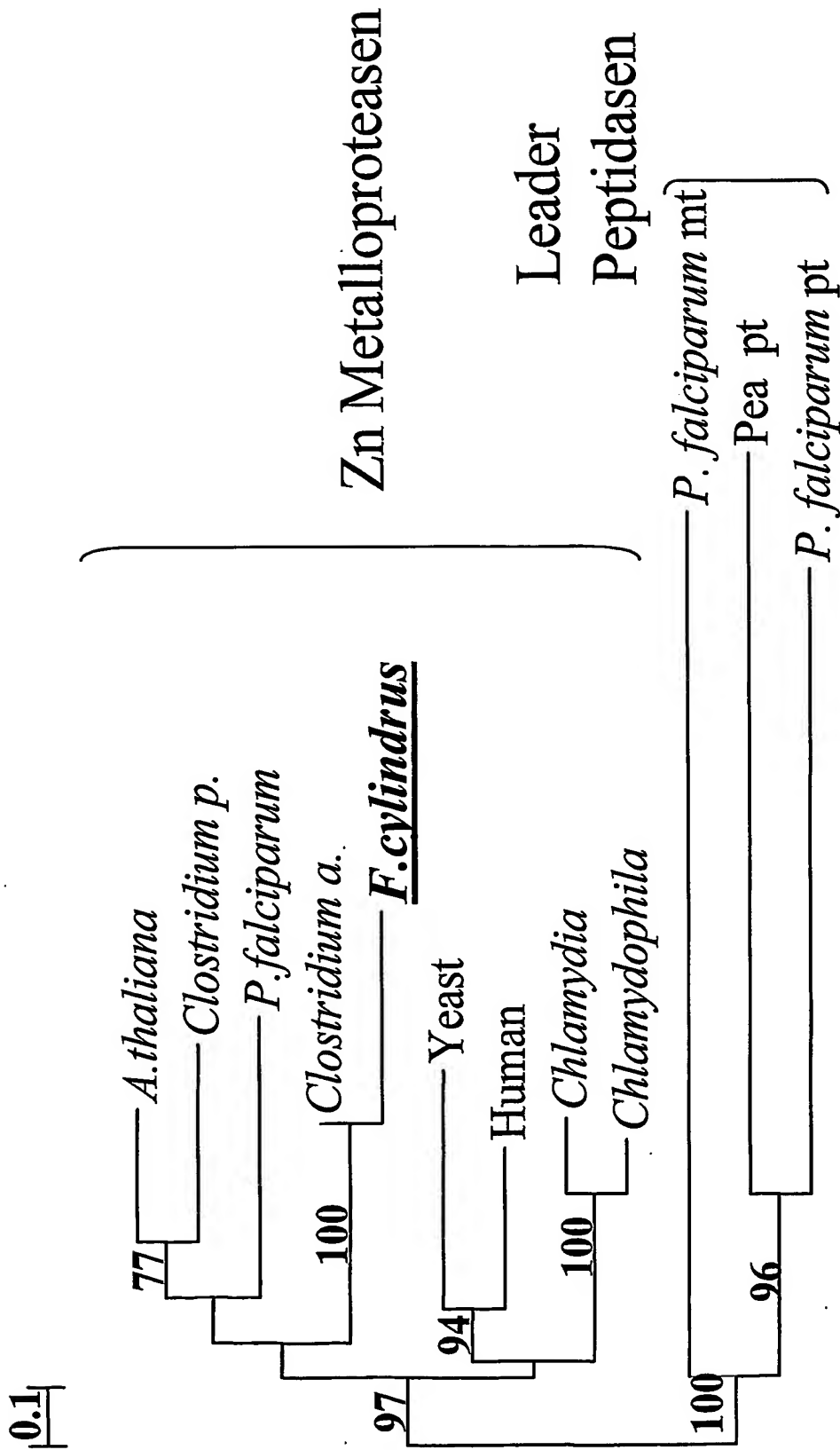


Fig.2

Sequenzprotokoll

5 <110>Stiftung Alfred-Wegener-Institut für Polar- und Meeresforschung,
 Bremerhaven
 <120>Für eine Calpain-Protease kodierende neue Nukleinsäuresequenz aus
 einer kälteangepassten marinen Diatomee *Fragilariopsis cylindrus*
 10 <130>AWI 01/0902 DE
 <160>2
 <210>1
 <211>544
 15 <212>DNA
 <213>*Fragilariopsis cylindrus*
 <400>1
 20 GG GAA TTC GGC CTT ACG GCC GGG GAT GAT GGA ATG TTC TGG ATT AGT 47
 Glu Phe Gly Leu Thr Ala Gly Asp Asp Gly Met Phe Trp Ile Ser 15
 1
 5
 10
 25 TGG GAG GAT GTC TTG CTT TAT TTC CGC AAT TTA CAA TTA TCA TGG AAT 95
 Trp Glu Asp Val Leu Tyr Phe Arg Asn Leu Gln Leu Ser Trp Asn 30
 20
 25
 30 CCC AAA CTA TTT GCG TAT CGG ATG ACT ACT CAT GGC TTA TGG CCA AAG 143
 Pro Lys Leu Phe Ala Tyr Arg Met Thr Thr His Gly Leu Trp Pro Lys 45
 35
 35 GAT CAG GGA CCA CAA AAT GAT GCA TTT AAT GTC GGA GAG AAT CCA CAA 191
 Asp Gln Gly Pro Gln Asn Asp Ala Phe Asn Val Gly Glu Asn Pro Gln 60
 50
 35
 40 TAT ATC ATG TCT TTC TCC GAA AAA GCT GTA TCG AGT AAA CCA ACG ATT 239
 Tyr Ile Met Ser Phe Ser Glu Lys Ala Val Ser Ser Lys Pro Thr Ile 75
 65
 70
 40
 40 TGG GTA CTG ATA TCA AGG CAT GTA AGC AAA CAG GAG CAA GAA GGT GCT 287
 Trp Val Leu Ile Ser Arg His Val Ser Lys Gln Glu Gln Glu Gly Ala 95
 80
 85
 90
 45
 45 GAG GTG AAT GAT TTC TTA ACC ATA CAT CTC GTT AGA AAC TCG GCT ACA 335
 Glu Val Asn Asp Phe Leu Thr Ile His Leu Val Arg Asn Ser Ala Thr 110
 100
 105
 110
 50
 50 TTA GAA AGA GTT TGG TAT CCC CAT GGA AAA GCA ACG ATT GCT AAT GGA 383
 Leu Glu Arg Val Trp Tyr Pro His Gly Lys Ala Thr Ile Ala Asn Gly 125
 115
 120
 125
 55
 55 TGC TAT ACA AAC AAT CCA CAC GTG CTT TTA CGA TAC GAT GTT TCC GGA 431
 Cys Tyr Thr Asn Asn Pro His Val Leu Leu Arg Tyr Asp Val Ser Gly 140
 130
 135
 140
 60
 60 CCT GAA GAT CAA TTT ATC TCG TTA GTA CTG TCT CAA CAC GAA AAA ACT 479
 Pro Glu Asp Gln Phe Ile Ser Leu Val Leu Ser Gln His Glu Lys Thr 155
 145
 150
 155
 60
 60 CAA GAT CTA TCA TAC ACT CTC TCT TGT TAC TGT ACC GAA CCC TTT ACA 527
 Gln Asp Leu Ser Tyr Thr Leu Ser Cys Tyr Cys Thr Glu Pro Phe Thr

| | | | | | | | | |
|----|---|--|-----|--|-----|--|-----|-----|
| | 160 | | 165 | | 170 | | 175 | |
| | CTA GGA AGA CCA CCA AA | | | | | | | 544 |
| 5 | Leu Gly Arg Pro Pro | | 180 | | | | | |
| | <210>2 | | | | | | | |
| | <211>544 | | | | | | | |
| | <212>DNA | | | | | | | |
| 10 | <213>Fragilariaopsis cylindrus | | | | | | | |
| | <400>2 | | | | | | | |
| 15 | TCA AAC GAT GGT GCG CAA TAC GTA GTA GAG AAA TCG ATA CTG GTA GGT | | | | | | | 48 |
| | Ser Asn Asp Gly Ala Gln Tyr Val Val Glu Lys Ser Ile Leu Val Gly | | | | | | | |
| | 1 .5 10 15 | | | | | | | |
| 20 | TCA GTG AAT TAT CCT GTA AAA GAT CCA TTT AAT CAG ATG AAA CGT GGA | | | | | | | 96 |
| | Ser Val Asn Tyr Pro Val Lys Asp Pro Phe Asn Gln Met Lys Arg Gly | | | | | | | |
| | 20 25 30 | | | | | | | |
| 25 | TCA CTT CAA ACC TAC TCA GAT TCA TGG ACC GAA CGG GAT CGT ACC TCA | | | | | | | 144 |
| | Ser Leu Gln Thr Tyr Ser Asp Ser Trp Thr Glu Arg Asp Arg Thr Ser | | | | | | | |
| | 35 40 45 | | | | | | | |
| 30 | TTT GTC ATG GCA TCA CGT AAC TTA GCC GAT TTT CGT AAT AAC GTG AAG | | | | | | | 192 |
| | Phe Val Met Ala Ser Arg Asn Leu Ala Asp Phe Arg Asn Asn Val Lys | | | | | | | |
| | 50 55 60 | | | | | | | |
| 35 | GTA ACG ATC GAT GCT GTT TTT AAT CCA CTT TTT ATC AAC GAG GAA TAC | | | | | | | 240 |
| | Val Thr Ile Asp Ala Val Phe Asn Pro Leu Phe Ile Asn Glu Glu Tyr | | | | | | | |
| | 65 70 75 80 | | | | | | | |
| 40 | AAA TGG ATC TTT CGT CAA GAA GGC TGG AGG TTA GAG ACA CCT GAC AAT | | | | | | | 288 |
| | Lys Trp Ile Phe Arg Gln Glu Gly Trp Arg Leu Glu Thr Pro Asp Asn | | | | | | | |
| | 85 90 95 | | | | | | | |
| 45 | GTC AAC CTA CTT ATC AAT GGG AAC GCT TAT GTA AAC GCT AAG GCC GAC | | | | | | | 336 |
| | Val Asn Leu Leu Ile Asn Gly Asn Ala Tyr Val Asn Ala Lys Ala Asp | | | | | | | |
| | 100 105 110 | | | | | | | |
| 50 | CAG ATG GAC CCC CAA GAG GTT ATG ATA AAG CAA ATC TAC AGC AAT CTC | | | | | | | 384 |
| | Gln Met Asp Pro Gln Glu Val Met Ile Lys Gln Ile Tyr Ser Asn Leu | | | | | | | |
| | 115 120 125 | | | | | | | |
| 55 | TTT GCT GAT CAC GTG TAT AGC AAA AGT CCA AAA GGA GAC GCC GCC CAA | | | | | | | 432 |
| | Phe Ala Asp His Val Tyr Ser Lys Ser Pro Lys Gly Asp Ala Ala Gln | | | | | | | |
| | 130 135 140 | | | | | | | |
| 60 | GTA GTC ACC ATG ACA TTG GCA CCA AGG GCG AAT TCT GCA GAT ATC CAT | | | | | | | 480 |
| | Val Val Thr Met Thr Leu Ala Pro Arg Ala Asn Ser Ala Asp Ile His | | | | | | | |
| | 145 150 155 160 | | | | | | | |
| 65 | CAC ACT GGC GGC CGT CTC GAG CAT GCA TCT AGA GGG CCC AAT TCG CCC | | | | | | | 528 |
| | His Thr Gly Gly Arg Leu Glu His Ala Ser Arg Gly Pro Asn Ser Pro | | | | | | | |
| | 165 170 175 | | | | | | | |
| 70 | TAT AGT GAG TCG TAT T | | | | | | | 544 |
| | Tyr Ser Glu Ser Tyr | | | | | | | |
| | 180 181 | | | | | | | |